

THEOREMS CONCERNING PROBABILITY



By

WILLIAM DOWELL BATEN

Amherst College

1888

THEOREMS CONCERNING PROBABILITY

A DISSERTATION

*Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in
the University of Michigan*

By

WILLIAM DOWELL BATEN

*20200



C O N T E N T S

CHAPTER I

	PAGE
Frequency laws for the sum of n variables which are subject each to given frequency laws	1

CHAPTER II

The probable error of certain functions of the errors made in measurements	17
--	----

CHAPTER III

Comparison of certain probabilities	29
-------------------------------------	----

CHAPTER IV

The evaluation of certain definite integrals by the use of probability functions	35
--	----

CHAPTER V

Simultaneous treatment of discrete and continuous probability by use of Stieltjes Integrals	43
---	----



Digitized by the Internet Archive
in 2026

CHAPTER I

Frequency Laws for the Sum of n Variables which are Subject each to given Frequency Laws *

INTRODUCTION.

It is the object of this article to find frequency laws for the sum of n independent variables, which are continuous and which have known continuous frequency laws. Laws for the sum of the variables will be found by use of results obtained by DODD (1).

Considerable work had been done along this phase of probability. CZUBER (2), using the Gaussian law for the individual variables obtained the law for the sum of the variables and also a law for the sum of the squares of the variables. MAYR (3) secured frequency laws for the sum, but in many cases restricted the individual variables to take on only positive values. $F(x_i) = 1/d$, where d is a constant, was used frequently. MISES discovered frequency laws for the sum of the variables, where the number of variables is taken very large, in fact approaches infinity (4). These are of course approximations. DODD also found probability functions for the asymptotic case (5); and used laws for the sum in comparing the arithmetic

(1) *The Frequency Law of a Function of one Variable*, by DODD. "Am. Math. Sc. Bul." 31, (1925), pages 27-31.

The Frequency Law of a Function of Variables with given Frequency Laws, by DODD. "An. of Math." 2 ser. vol. 27, No. 1.

(2) *Theorie der Beobachtungsfehler*, by CZUBER, 1891, pages 72-150.

(3) *Wahrscheinlichkeitsfunktionen und ihre Anwendungen*, by MAYR. "Monatshefte für Math. und Phys." Vol. 30.

(4) *Fundamentalsätze der Wahrscheinlichkeitsrechnung*, by MISES. "Math. Zeitschrift" 1919, Vol. 4, pages 1-97.

(5) *Functions of Measurements under General Laws of Error*, by DODD. "Sarttryck ur Skandinavisk Aktuarietidskrift". 1922.

* Published in Metron Vol. X - N. 3, 1932.

mean with other functions of the measurements, such as the geometric mean, root-mean-square and the median (6). POINCARÉ derived laws for the sum by using his characteristic functions (7).

GENERAL METHOD.

Let x_i be subject to the frequency law $f_i(x_i)$ which is defined in some interval and whose definite integral over this interval is unity. In many cases considered this interval is $(-\infty, \infty)$. Each variable x_i , ($i = 1, 2, 3, \dots, n$), is subject to the same law for the particular problem. The following theorem was used for determining the probability or frequency law for the sum (1).

Let $f_i(x_i)$ be the frequency function for x_i , where $i = 1, 2, \dots, n$, these functions being non-negative and integrable from minus infinity to plus infinity with result unity. Set

$$C_i = \int_{-\infty}^{\infty} f_i(x_i) \cos(t x_i) dx_i ; S_i = \int_{-\infty}^{\infty} f_i(x_i) \sin(t x_i) dx_i ;$$

$$r_i = \sqrt{C_i^2 + S_i^2} ; \theta_i = \arcsin(S_i/r_i) = \arccos(C_i/r_i) ;$$

$$R = \prod_{i=1}^n r_i ; \theta = \sum_{i=1}^n \theta_i.$$

Let $f_1(x_1)$ be of limited variation, and let each f be either continuous or of limited variation. Then the probability $P(u)$ that

$$x_1 + x_2 + x_3 \dots + x_n \leq u,$$

is

$$P(u) = 1/2 + \frac{1}{\pi} \int_0^{\infty} R \sin(ut - \theta) dt/t.$$

If furthermore $f_2(x_2)$ is of limited variation

$$p(u) = dP/du = \frac{1}{\pi} \int_0^{\infty} R \cos(ut - \theta) dt.$$

The detailed work will be given for two cases, while results for

(6) *The Probability of the Arith. Mean compared with that of other Functions of the Measurements*, by DODD. "An. of Math.", Vol. 14, 1913, pages 186-198.

(7) *Calcul des Probabilités*, by POINCARÉ, 1912.

other cases will be exhibited in a table. $p(u)$ will be used instead of $P(u)$.

Case I.

Let $f_i(x_i) = f_2(x_2) = \dots = f_n(x_n) = \frac{2h}{\pi} \cdot \frac{1}{(1 + h^2 x_i^2)^2}$, where $f_i(x_i) dx_i$ is the probability that x_i lies in $(x_i, x_i + dx_i)$. The area under the curve $f_i(x_i)$ between x_i and $x_i + dx_i$ represents the probability that x_i lies in this interval.

Each variable has its own probability law. Here all of the laws are the same. The problem here is to find a function which is non-negative such that the area under the graph of this function between u and $u + du$ will represent the probability that the sum

$$u \leq \sum_{i=1}^n x_i \leq u + du.$$

From the theorem

$$C_1 = C_2 = \dots = C_n = \frac{2h}{\pi} \int_{-\infty}^{\infty} \frac{\cos(tx_i) dx_i}{(1 + h^2 x_i^2)^2} = \frac{h+t}{h} \cdot e^{-t/h};$$

$$S_1 = S_2 = \dots = S_n = \frac{2h}{\pi} \int_{-\infty}^{\infty} \frac{\sin(tx_i) dx_i}{(1 + h^2 x_i^2)^2} = 0;$$

$$r_1 = r_2 = \dots = r_n = \frac{(h+t)e^{-t/h}}{h}; \quad \theta_i = 0;$$

$$R = C_i^n = \frac{(h+t)^n e^{-n t/h}}{h^n}, \quad \theta = 0.$$

$$p_n(u) = \frac{1}{h^n \pi} \int_0^{\infty} (h+t)^n e^{-n t/h} \cos(ut) dt,$$

or

$$p_n(u) = \frac{1}{h^n \pi} \sum_{r=1}^{n+1} \binom{n}{r-1} \frac{h^{n+1} \Gamma(r) \cos(r \cdot \arctan hu/n)}{(n^2 + h^2 u^2)^{r/2}},$$

which is the required probability.

If n is equal to 2 and 3, $p_n(u)$ becomes

$$p_2(u) = \frac{4h(20 + h^2 u^2)}{\pi(2^2 + h^2 u^2)^3},$$

and

$$p_3(u) = \frac{6h(1053 + 54h^2u^2 + h^4u^4)}{\pi(3^2 + h^2u^2)^4}.$$

These prove to be probability functions for they are non-negative with integral from $-\infty$ to $+\infty$ equal to unity. The interval here is $(-\infty, +\infty)$ because the sum of the variables may have any real value.

Case 2.

Let $f_i(x_i) = (h/4 + (h^2/4)|x_i|)e^{-h|x_i|}$ be the probability that x_i , $i = 1, 2, \dots, n$, lies in the interval $(x_i, x_i + dx_i)$; find the probability that

$$u \leq x_1 + x_2 \dots + x_n \leq u + du.$$

From the theorem

$$C_1 = C_2 \dots = C_n = \int_{-\infty}^{\infty} (h/4 + (h^2/4)|x_i|) \cos(tx_i) dx_i = \frac{h^4}{(h^2 + t^2)^2};$$

$$S_1 = S_2 = \dots = S_n = 0; \theta_1 = \theta_2 = \dots \theta_n = 0;$$

$$r_1 = r_2 = \dots = r_n = \frac{h^4}{(h^2 + t^2)^2}; \theta = \sum_{i=1}^n \theta_i = 0;$$

$$R = \frac{h^{4n}}{(h^2 + t^2)^{2n}}$$

Then

$$p_n(u) = \frac{h^{4n}}{\pi} \int_0^{\infty} \frac{\cos(tu) dt}{(h^2 + t^2)^{2n}},$$

which is the probability that $u \leq \sum_{i=1}^n x_i \leq u + du$.

For different values of n the above integral can be integrated. For n equal 2 and 3

$$p_2(u) = \frac{h}{96} (15 + 15h|u| + 6h^2u^2 + h^3|u|^3) e^{-h|u|},$$

$$p_3(u) = \frac{h}{5!2^6} (945 + 945h|u| + 420h^2u^2 + 105h^3|u|^3 + \\ + 15h^4u^4 + h^5|u|^5) e^{-h|u|}.$$

EVALUATION OF A USEFUL INTEGRAL.

Since several probability functions come out in the form of an integral similar to $\int_0^{\infty} \frac{\cos(ux) dx}{(h^2 + x^2)^n}$, it seems worth while to find a general law for integrating such integrals. This can be changed to the form $\int_0^{\infty} \frac{\cos(ux) dx}{(1 + x^2)^n}$. Integrate $\frac{e^{iuz}}{(1 + z^2)^n} = f(z)$ around the contour C , which encloses the singular point $z = i(1)$. The residue of $f(z)$ at this singular point is the value of $\frac{1}{2\pi i} \int_C f(z) dz$; this residue is also equal to the coefficient of z^{-1} in the Laurent expansion of $f(z)$ in a power series. The residue of $f(z)$ at i is the same as the residue of $f(i + w)$ at $w = 0$.

$$\begin{aligned}
 f(i + w) &= \frac{e^{-u} e^{iuz}}{(1 - 1 + 2iw + w^2)^n} = \frac{e^{-u} e^{iuz}}{2^n w^n i^n \left(1 - \frac{iw}{2}\right)^n} \\
 &= \frac{e^{-u} e^{iuz} \left(1 - \frac{iw}{2}\right)^{-n}}{2^n w^n i^n} = \frac{e^{-u}}{2^n w^n i^n} \left\{ 1 + iuw - \frac{u^2 w^2}{2!} - \right. \\
 &\quad \left. - \frac{i u^3 w^3}{3!} + \frac{u^4 w^4}{4!} \dots \right\} \left\{ 1 + \binom{-n}{1} \left(\frac{-iw}{2}\right) + \binom{-n}{2} \left(\frac{-iw}{2}\right)^2 \dots \right\} \\
 &= \frac{e^{-u} w^{n-1}}{2^n w^n i^n} \left\{ \frac{\binom{-n}{n-1} (-i)^{n-1}}{2^{n-1} 0!} + \frac{\binom{-n}{n-2} (-i)^{n-2} iu}{2^{n-2} 1!} + \right. \\
 &\quad \left. + \frac{\binom{-n}{n-3} (-i)^{n-3} i^2 u^2}{2^{n-3} 2!} + \dots + \frac{\binom{-n}{2} (+i)^{n-3} (-i)^2 u^{n-3}}{2^2 (n-3)!} + \right. \\
 &\quad \left. + \frac{\binom{-n}{1} (-i) i^{n-2} u^{n-2}}{2^1 (n-2)!} + \frac{\binom{-n}{0} u^{n-1} i^{n-1}}{2^0 (n-1)!} \right\} + \dots
 \end{aligned}$$

(1) The contour C consists of x -axis from $-R$ to $+R$, the upper half of the circle where $|z| = R$. R is allowed to approach infinity.

The terms involving w^{-1} are only considered. This reduces to

$$\frac{e^{-u}}{2^n i} \left\{ \frac{n^{(-n-1)} u^0}{2^{n-1} 0! (n-1)!} + \frac{n^{(-n-2)} u^1}{2^{n-2} 1! (n-2)!} + \frac{n^{(-n-3)} u^2}{2^{n-3} 2! (n-3)!} + \dots \right. \\ \left. + \frac{n^{(-2)} u^{n-3}}{2^2 (n-2)! 2!} + \frac{n^{(-1)} u^{n-2}}{2^1 (n-2)! 1!} + \frac{n^{(0)} u^{n-1}}{2^0 (n-1)! 0!} \right\} w^{-1}.$$

The residue of $f(i+w)$ at $w=0$ is the coefficient of w^{-1} , which reduces to

$$\frac{e^{-u}}{2^n i} \sum_{r=0}^{n-1} \frac{n^{(-r)} u^{n-1-r}}{2^r (n-1-r)! r!}, \text{ where } n > 0. \quad (1)$$

The residue is also the value of $\frac{1}{2\pi i} \int_C f(z) dz$; hence

$$\frac{1}{2\pi i} \int_C f(z) dz = \frac{e^{-u}}{2^n i} \sum_{r=0}^{n-1} \frac{n^{(-r)} u^{n-1-r}}{2^r (n-1-r)! r!},$$

$$\frac{1}{\pi} \int_0^\infty \frac{\cos(ux) dx}{(1+x^2)^n} = \frac{e^{-u}}{2^n} \sum_{r=0}^{n-1} \frac{n^{(-r)} u^{n-1-r}}{2^r (n-1-r)! r!}, \text{ where } u > 0.$$

This integral was evaluated by B. DE HAAN AS an infinite series. The above evaluation contains only a finite number of terms. This integral was evaluated by another method by SERRET (2).

Since the above integral has been found, the probability for the sum of the n variates, mentioned in Case 2, can be found at once for any n .

Case 3.

Let it be required to find the probability that

$$u \leq \sum_{i=1}^n x_i \leq u + du,$$

where the probability for x_i is $f_i(x_i) = \frac{h}{2} e^{-h|x_i|}$. According to the

(1) $n^{(-r)} = (n)(n+1)(n+2) \dots (n+r-1)$, $n^{(0)} = 1$.

(2) *Cours de Calcul Différentiel et Intégral*, by SERRET. Vol. 2, 1894, pages 199, 131.

theorem

$$p_n(u) = \frac{h^{2n}}{\pi} \int_0^\infty \frac{\cos(ut) dt}{(h^2 + t^2)^n} = \frac{1}{\pi} \int_0^\infty \frac{\cos(ut) dt}{\left\{1 + \left(\frac{t}{h}\right)^2\right\}^n}.$$

Let $hy = t$, then $dt = h dy$ and

$$p_n(u) = \frac{h}{\pi} \int_0^\infty \frac{\cos(huy) dy}{(1 + y^2)^n}.$$

According to the above general formula for integrating this kind of integral

$$p_n(u) = \frac{h^{n-1}}{2^n} \sum_{r=0}^{n-1} \frac{n^{(-r)} |hu|^{n-1-r}}{2^r (n-1-r)! r!} e^{-k|u|}.$$

It is necessary to introduce the absolute value signs since the general formula for evaluating this integral is true when $u > 0$.

MAYR (3) considered this case and found a long process of discovering the law for the sum, but gave no general formula for any n . The above method, by using the integral developed, gives results with little difficulty.

If to each variable x_i the constant k is added and each variable is subject to the Gaussian law $\frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$, then the law for the sum

$$(x_1 + k) + (x_2 + k) + \dots + (x_n + k)$$

is

$$p_n(u) = \frac{h}{\sqrt{n\pi}} e^{-h^2(u-nk)^2/n}.$$

LAW FOR THE SUM OF THE SQUARES OF THE VARIABLES.

Let the individual variables be subject to the Gaussian law $\frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$, with precision h , and let it be required to find the probability that

$$u \leq x_1^2 + x_2^2 + \dots + x_n^2 \leq u + du.$$

Using the theorem mentioned above

$$C_i = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-h^2 x_i^2} \cos(t x_i^2) dx_i = \frac{\cos\left(\frac{1}{2} \arctan t/h^2\right)}{h^{-1}(h^4 + t^2)^{1/4}};$$

$$S_i = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-h^2 x_i^2} \sin(t x_i^2) dx_i = \frac{\sin\left(\frac{1}{2} \arctan t/h^2\right)}{h^{-1}(h^4 + t^2)^{1/4}};$$

$$r_i = \frac{h}{(h^4 + t^2)^{1/4}}, \quad \theta_i = (1/2) \arctan t/h^2 = \beta;$$

$$R = \prod_{i=1}^n r_i = \frac{h^n}{(h^4 + t^2)^{n/4}}, \quad \theta = n\beta;$$

$$p_n(u) = \frac{h^n}{\pi} \int_0^{\infty} \frac{\cos(ut - nm)}{(h^4 + t^2)^{n/4}} dt, \quad \text{where } 2m = \arctan t/h^2.$$

$$p_n(u) = \frac{h^n}{\Gamma(n/2)} u^{n/2-1} e^{-h^2 u}.$$

This shows that the frequency law for the sum of the squares is not a Gaussian law when the individual variables are subject to the Gaussian law. It is a Type 3 frequency law.

MEASURE OF PRECISION.

In the Gaussian law $f_i(x_i) = \frac{h_i}{\sqrt{\pi}} e^{-h_i^2 x_i^2}$, h_i is considered to be the "measure of precision". Since the mean deviation is $\frac{1}{h_i \sqrt{\pi}}$, and the second moment about the mean is $1/2 h_i^2$, h_i is considered as the measure of precision for the larger h_i is, the smaller the dispersion and standard deviation. If the variables x_i , ($i = 1, 2, \dots, n$), are each subject to the Gaussian law with measure of precision h_i , then the sum of the variables is also subject to the Gaussian law with measure of precision H , where

$$\frac{1}{H^2} = \frac{1}{h_1^2} + \frac{1}{h_2^2} + \frac{1}{h_3^2} + + + + + \frac{1}{h_n^2}.$$

Proof: According to the theorem which is used for finding the law for the sum

$$p_n(u) = \frac{1}{\sqrt{\pi \sum \frac{1}{h_i^2}}} e^{-\frac{u^2}{\sum 1/h_i^2}} = \frac{H}{\sqrt{\pi}} e^{-H^2 u^2}$$

This has been proven by other, but this method seems to be the simplest.

According to this method of finding the probability law for the sum, the law for the weighted mean

$$M = p_1 x_1 + p_2 x_2 + p_3 x_3 + \dots + p_n x_n,$$

where $\sum p_i = 1$, is

$$p_n(u) = \frac{H}{\sqrt{\pi}} e^{-H^2 u^2}, \text{ where } \frac{1}{H} = \sqrt{\sum_{i=1}^n \frac{p_i^2}{h_i^2}},$$

provided the individual variables are subject to the Gaussian law with precision h_i . In this case the law for the quantity $p_i x_i$ must

be found. If the probability function for x_i is $f_i(x_i) = \frac{h_i}{\sqrt{\pi}} e^{-h_i^2 x_i^2}$,

the probability function for $p_i x_i$ is, according to reference (1); F_i

$$(p_i x_i) = \frac{h_i}{p_i \sqrt{\pi}} e^{-(h_i/p_i)^2 x_i^2}$$

If the frequency law of the individual variable is

$$f_i(x_i) = \frac{h_i}{\pi} \frac{1}{1 + h_i^2 x_i^2}, \text{ where } h_i \text{ enters the function as a variable}$$

parameter, then the variable parameter which occurs in the frequency law for the sum of n variables is $1/H$, where

$$\frac{1}{H} = \frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{h_3} + \dots + \frac{1}{h_n}.$$

This is similar to the result obtained when the individual variables were subject to the Gaussian function.

If the variable x_i is subject to the law $f_i(x_i) = \frac{h_i}{2} e^{-h_i |x_i|}$, where

h_i is the variable parameter, then the law for the sum of n independent variables is

$$p_n(u) = \frac{\prod_{i=1}^n h_i^2}{\pi} \int_0^{\infty} \frac{\cos(ut) dt}{\prod_{i=1}^n (h_i^2 + t^2)}.$$

If this is broken up into partial fractions and integrated it becomes

$$p_n(u) = \frac{K}{2\pi} \sum_{i=1}^n Q_i e^{-h_i |u|},$$

where

$$K = \prod_{i=1}^n h_i^2, \quad Q_i = \frac{1}{\prod_{s=1}^{i-1} (h_s^2 - h_i^2) \prod_{s=i+1}^{n-1} (h_i^2 - h_s^2) h_i},$$

where $i + s < n - 1$, and any parenthesis with $h_0^2 = 1$. It is assumed that all of the h_i 's are distinct. If several were equal the fraction would be broken up into different partial fractions and of course $p_n(u)$ would be different. Here the variable parameter enters into the law for the sum in a rather complicated way.

CHARLIER DEVELOPMENT OF CERTAIN FREQUENCY FUNCTIONS.

Case I.

The frequency function for the sum of the squares of n independent variables, when each variable is subject to the Gaussian function can be exhibited in a manner similar to a Charlier series. Consider this function in the integral form.

$$p_n(u) = \frac{h^n}{\pi} \int_0^{\infty} \frac{\cos(ut - nm) dt}{(h^4 + t^2)^{n/4}}. \quad (\text{See page 7}).$$

$$p_2(u) = h^2 \left\{ \frac{h^2}{\pi} \int_0^{\infty} \frac{\cos(ut) dt}{(h^4 + t^2)} + \frac{1}{\pi} \int_0^{\infty} \frac{t \sin(ut) dt}{(h^4 + t^2)} \right\}.$$

Let

$$F_1(u) = \frac{1}{\pi} \int_0^{\infty} \frac{\cos(ut) dt}{(h^4 + t^2)^i},$$

then

$$p_2(u) = \{h^2 F_1(u) - F_1'(u)\} h^2 = h^2(h^2 - D) F_1(u),$$

where D is an operator and represents differentiation.

$$p_4(u) = \left\{ \frac{h^4}{\pi} \int_0^{\infty} \frac{\cos(ut) dt}{(h^4 + t^2)^2} + \int_0^{\infty} \frac{2h^2 t \sin(ut) dt}{\pi(h^4 + t^2)^2} - \frac{1}{\pi} \int_0^{\infty} \frac{t^2 \cos(ut) dt}{(h^4 + t^2)^2} \right\} h^4,$$

$$p_4(u) = \{h^4 F_2(u) + 2F_2'(u)h^2 - F_2''(u)\} h^4 = h^4(h^2 - D)^2 F_2(u).$$

Similarly

$$p_{2n}(u) = h^{2n}(h^2 - D)^n F_n(u).$$

The operator D raised to a power k represents the k -th derivative. This is similar to a Charlier series since the function is expressed as the sum of the first term and its derivatives. Here the series is finite and the coefficients are coefficients in the binomial expansion. It must be noted that the above is only for values of n when it is an even integer.

$$\text{Let } G_i(u) = \frac{1}{\pi} \int_0^{\infty} \frac{\cos(ut) (h^4 + t^2 + h^2)^{1/2} dt}{2(h^4 + t^2)^{i/4}}, \quad i = \text{odd integer},$$

$$Z_i(u) = \frac{1}{\pi} \int_0^{\infty} \frac{\sin(ut) (h^4 + t^2 - h^2)^{1/2} dt}{2(h^4 + t^2)^{i/4}}$$

then

$$p_{2n+1}(u) = \{(h^2 - D)^{n+1} G_{2n+1}(u) + (h^2 - D)^{n+1} Z_{2n+1}(u)\} h^{2n+1}.$$

$p_n(u)$ with odd integer subscript is the sum of two series.

When each variable is subject to the frequency function $h^2 e^{-h^2 x_i} x_i$, which is defined over $(0, \infty)$, then the frequency function for the sum of n independent variables is

$$p_n(u) = \frac{h^{2n}}{\pi} \int_0^{\infty} \frac{\cos(ut - n\theta) dt}{(h^2 + t^2)^n}, \quad \cos \theta = \frac{h^2 - t^2}{h^2 + t^2},$$

$$p_2(u) = \frac{h^4}{\pi} \int_0^{\infty} \frac{(h^4 - 6h^2t^2 + t^2) \cos(ut) dt}{(h^2 + t^2)^4} + \frac{h^4}{\pi} \int_0^{\infty} \frac{(4h^3t - 4ht^3) \sin(ut) dt}{(h^2 + t^2)^4}.$$

Let

$$F_4(u) = \frac{1}{\pi} \int_0^{\infty} \frac{\cos(ut) dt}{(h^2 + t^2)^4},$$

then

$$p_2(u) = h^4 (h - D)^4 F_4(u).$$

$$p_n(u) = h^{2n} (h - D)^{2n} F_{2n}(u);$$

which is similar to a Charlier series, because each term after the first is a function of the derivatives of this term.

If the law for the individual variables is $\frac{h^r}{\Gamma(r)} e^{-hx} x^{r-1}$, where $r > -1$, then the law for the sum of n variables is

$$p_n(u) = \frac{h^{nr}}{\pi} \int_0^{\infty} \frac{\cos(ut - nrk) dt}{(h^2 + t^2)^{nr/2}}, \text{ where } k = \arctan h.$$

For different values of r and n this can be expressed in a Charlier series where the series contains only a finite number of terms. The individual variables are defined for this last case in the interval $(0, \infty)$.

It seems of interest to be able to express certain probability functions which have only a finite number of terms into a series which has also a finite number of terms, where each term after the first term is equal to a determined coefficient multiplied by a derivative of the first term. The last case given suggests that there are many such probability functions.

EVALUATION OF USEFUL INTEGRALS.

Since

$$\frac{1}{\pi} \int_0^{\infty} \frac{\cos(ux) dx}{(1+x^2)^n} = \sum_{r=0}^{n-1} \frac{n^{(-r)} u^{n-1-r} e^{-u}}{2^r (n-1-r)! (r)!},$$

the value of integrals

$$(A) \quad \frac{1}{\pi} \int_0^{\infty} \frac{x^k \sin(ux) dx}{(1+x^2)^n}, \quad \frac{1}{\pi} \int_0^{\infty} \frac{x^k \cos(ux) dx}{(1+x^2)^n}$$

can be found by differentiating the above with respect to u . Let

$$p_n(u) = \frac{1}{\pi} \int_0^{\infty} \frac{\cos(ux) dx}{(1+x^2)^n}, \quad q_n(u) = \frac{1}{\pi} \int_0^{\infty} \frac{x \sin(ux) dx}{(1+x^2)^n},$$

then

$$(2) \quad q_{n+1}(u) = \frac{u}{2n} p_n(u).$$

Hence it is not difficult to evaluate $q_n(u)$.

Differentiating (2) with respect to u

$$q'_{n+1}(u) = \frac{p_n(u)}{2n} + \frac{p'_n(u) \cdot u}{2n}, \text{ or}$$

$$\frac{1}{\pi} \int_0^{\infty} \frac{x^2 \cos(ux) dx}{(1+x^2)^{n+1}} = \frac{1}{2n\pi} \int_0^{\infty} \frac{\cos(ux) dx}{(1+x^2)^n} - \int_0^{\infty} \frac{u x \sin(ux) dx}{2n\pi (1+x^2)^n},$$

which makes possible the evaluation of the left member since both terms in the right member are now known.

By successive differentiation of (2) the integrals (A) can be evaluated. It is understood that k must not be so that the integrals do not exist. k must be > -2 in fact, $k \neq 2n$.

The above integrals appear in finding frequency laws for functions of the n variables, in the Dodd method. In the law for the sum of the squares they must be evaluated.

Let

$$p_n(u) = \frac{1}{\pi} \int_0^{\infty} \frac{\cos(ux) dx}{(1+x^2)^n}, \quad q_n(u) = \frac{1}{\pi} \int_0^{\infty} \frac{x \sin(ux) dx}{(1+x^2)^n};$$

$$p_3(u) = \frac{1}{2!2^3} (3 + 3u + u^2) e^{-u},$$

$$p_4(u) = \frac{1}{3!2^4} (15 + 15u + 6u^2 + u^3) e^{-u},$$

$$p_3(u) = \frac{1}{4!2^5} (105 + 105u + 45u^2 + 10u^3 + u^4) e^{-u},$$

$$p_6(u) = \frac{1}{5!2^6} (945 + 945u + 420u^2 + 105u^3 + 15u^4 + u^5) e^{-u},$$

$$p_7(u) = \frac{1}{6!2^7} (10395 + 10395u + 4725u^2 + 1260u^3 + 210u^4 + 21u^5 + u^6) e^{-u},$$

$$p_8(u) = \frac{1}{7!2^8} (145135 + 145135u + 72765u^2 + 17325u^3 + 3150u^4 + 368u^5 + 28u^6 + u^7) e^{-u},$$

$$q_4(u) = \frac{1}{3!2^4} (3u + 3u^2 + u^3) e^{-u},$$

$$q_5(u) = \frac{1}{4!2^5} (15u + 15u^2 + 6u^3 + u^4) e^{-u},$$

$$q_6(u) = \frac{1}{5!2^6} (105u + 105u^2 + 45u^3 + 10u^4 + u^5) e^{-u},$$

$$q_7(u) = \frac{1}{6!2^7} (945u + 945u^2 + 420u^3 + 105u^4 + 15u^5 + u^6) e^{-u},$$

$$q_8(u) = \frac{1}{7!2^8} (10395u + 10395u^2 + 4725u^3 + 1260u^4 + 210u^5 + 21u^6 + u^7) e^{-u},$$

$$q_9(u) = \frac{1}{8!2^9} (145135u + 145135u^2 + 72765u^3 + 17325u^4 + 3150u^5 + 368u^6 + 28u^7 + u^8) e^{-u},$$

$$-p_3^*(u) = \frac{1}{16} (1 + u - u^2) e^{-u},$$

$$-p_4^*(u) = \frac{1}{96} (3 + 3u - u^3) e^{-u},$$

$$-p_5^*(u) = \frac{1}{4!2^5} (15 + 15u + 3u^2 - 2u^3 - u^4) e^{-u}.$$

Frequency law for single variable $f_1(x_1)$	Interval	Frequency law for the sum of n variables $f_n(u)$	Frequency law for the mean of n variables $f_M(u)$
$\frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$	$(-\infty, \infty)$	$\frac{h}{\sqrt{\pi n}} e^{-h^2 u^2/n}$	$\frac{h\sqrt{n}}{\sqrt{\pi}} e^{-n h^2 u^2}$
$\frac{2h}{\pi (1 + h^2 x^2)^2}$	$(-\infty, \infty)$	$\frac{1}{h^n \pi} \sum_{r=1}^{n+1} \binom{n}{r-1} \frac{\Gamma(r) \cos(r \tan^{-1} h u/n)}{h - n - 1} (n^2 + h^2 u^2)^{r/2}$	Replace h by $n h$.
$\frac{2h}{\pi} \cdot \frac{1}{e^{h x} + e^{-h x}}$	$(-\infty, \infty)$	$\frac{2^n}{\pi} \int_0^\infty \frac{\cos(ut) dt}{(e^{\pi t/2 h} + e^{-\pi t/2 h})^n}$	$\frac{2^n}{\pi} \int_0^\infty \frac{\cos(ut) dt}{(e^{\pi t/2 n h} + e^{-\pi t/2 n h})^n}$
$\frac{h e^{-h^2(x^2 - 2ax)}}{\pi e^{h^2 a^2}}$	$(-\infty, \infty)$	$\frac{h}{\sqrt{n \pi}} e^{-h^2(u - a n)^2/n}$	$\frac{h\sqrt{n}}{\sqrt{\pi}} e^{-h^2 n(u - a n)^2}$
$\left(\frac{h}{4} + \frac{h^2}{4} x\right) e^{-h x }$	$(-\infty, \infty)$	$\frac{h^4 n}{\pi} \int_0^\infty \frac{\cos(ut) dt}{(h^2 + t^2)^{2n}}$	$\frac{(n h)^4 n}{\pi} \int_0^\infty \frac{\cos(ut) dt}{(n^2 h^2 + t^2)^{2n}}$
$\frac{h}{2} e^{-h x }$	$(-\infty, \infty)$	$h e^{-h x } \left \sum_{r=0}^{n-1} \frac{n(-r) h u ^{n-1-r}}{2^r (n-1-r)! (r)!} \right $	Replace h by $n h$.
$\frac{2 h^3}{\sqrt{\pi}} e^{-h^2 x^2}$	$(-\infty, \infty)$	$\frac{1}{\pi 2^n h^2 n} \int_0^\infty e^{-t^2 n/4 h^2 (2 h^2 - t^2)^n} \cos(ut) dt$	Replace h by $n h$.

Frequency law for single variable $f_1(x_1)$	Interval	Frequency law for the sum of n variables $p_n(u)$	Frequency law for the mean of n variables $p_M(u)$
$\frac{\Gamma(2k+2)(1-x^2)^k}{2^2 k+1(k+1)^2}, k=1, \frac{3}{4}, r$	$(-1, 1)$	$\frac{3^n}{\pi} \int_0^\infty \left\{ \frac{\sin t - t \cos t}{t^3} \right\}^n \cos(ut) dt$	
$\frac{h^r}{\Gamma(r)} e^{-h x_i} x_i^{r-1}, r > -1$	$(0, \infty)$	$\frac{h^n r}{\pi} \int_0^\infty \frac{\cos(ut - n r k) dt}{(h^2 + t^2)^{n r/2}}$	Replace h by $n h$.
$1/2 h$	$(-h, h)$	$\frac{1}{\pi h^n} \int_0^\infty \frac{(\sin h t)^n \cos(ut) dt}{t^n}$	Replace h by $n h$.
Type 3. $h(1+x/a)^a e^{-g x}$	$(-a, \infty)$	$\frac{a-n a g}{\pi} \frac{h^n \cos a g \Gamma(g a+1)^n}{\int_0^\infty \cos\{(u+n a)t - n(g a+1) \arctan t/g\} dt}$	
$\frac{h}{\pi} \cdot \frac{1}{1+h^2 x^2}$	$(-x, \infty)$	$\frac{n h}{\pi (u^2 h^2 + n^2)}$	$\frac{h}{\pi} \cdot \frac{1}{1+h^2 u^2}$, same for any n .
$\frac{2 h}{\pi} \cdot \frac{x^2}{(h^2 + x^2)^2}$	$(-\infty, \infty)$	$\sum_{r=1}^n \frac{(-1)^{r-1}}{(r-1)} \binom{n}{r-1} \frac{\Gamma(r) \cos(r \cdot \arctan u/nh)}{((n h)^2 + u^2)^{r/2}}$	
$\frac{\sqrt{h} e^{-h x}}{\sqrt{\pi \cdot x}}$	$(0, \infty)$	$\frac{h^{n/2}}{\pi} \int_0^\infty \frac{\cos(ut - n t) dt}{(h^2 + t^2)^{n/4}}, \tan 2k = t/h$	Replace h by $n h$.

CHAPTER II

THE PROBABLE ERROR OF CERTAIN FUNCTIONS
OF THE ERRORS MADE IN MEASUREMENTS*

1. *Introduction.* This paper presents with proof and applications several theorems pertaining to the probable error of certain functions of the errors made in measurements. If the measurements or observations are multiplied by a constant b , the probable value of functions of the error of bm will differ from the probable value of these same functions of the error of m the measurement. This article shows that there are values of b such that the probable value of certain functions of the error of bm is less than the probable value of these functions of the error of m .

The probable value of certain functions of the mean are also treated. General frequency laws for the errors are used in the first theorems; these include the discrete, the continuous and a combination of the discrete and the continuous cases. Here Stieltjes integrals are used in the proofs. Special cases are mentioned.

The following theorems are proven by use of general frequency laws which come under the continuous case. The Gaussian law is treated as special cases to the theorems.

2. *Concerning the Square of the Error of the Measurement.*

THEOREM 1. *Let x be the error of the measurement m , d the expected value of x , and c the expected value of x^2 . Then under any law of error, whose second moment with respect to the true value a exists, there exist values of the constant b such that the probable value of the square of the error of bm is less than the probable value of the square of the error of the measurement, provided*

$$ad - c \neq 0, \quad a^2 - 2ad + c \neq 0.$$

Under these conditions b lies between

$$1 \quad \text{and} \quad 1 - \frac{2(c - ad)}{a^2 - 2ad + c}.$$

* Presented to the American Mathematical Society, April 19, 1930. Published in the Bulletin of the American Mathematical Society, February, 1931.

PROOF. Let $F(x)$ be the probability that x lies in the interval $(-\infty, x)$ where x is included. Then the probability that x lies in $(-\infty, x^*)$, where x^* means that x has been deleted, is $F(x-0)$. The probability that x lies at x is $F(x+0) - F(x-0)$, the probability that x lies in the segment (x_1^*, x_2^*) is

$$F(x_2 - 0) - F(x_1 + 0) = \int_{x_1+0}^{x_2-0} dF(x).$$

Finally the probability that x lies in the closed interval (x_1, x_2) is

$$F(x_2) - F(x_1 - 0) = \int_{x_1-0}^{x_2} dF(x).$$

We have also $F(-\infty) = 0$ and $F(+\infty) = 1$. If $f(x)$ is the probability function and is continuous, then

$$(1) \quad F(x) = \int_{-\infty}^x f(x)dx.$$

The error of the measurement m from the true value a is $x = a - m$. The error of bm is

$$x' = a - bm = a - b(a - x) = a(1 - b) + bx.$$

The probable value of $(x')^2$ is

$$(2) \quad \begin{cases} P(x'^2) = \int_{-\infty}^{\infty} \{a^2(1-b)^2 + 2a(1-b)bx + b^2x^2\} dF(x) \\ = a^2(1-b)^2 + 2abd(1-b) + b^2c, \end{cases}$$

which is less than the probable value of the square of the error of m ,

$$P(x^2) = \int_{-\infty}^{\infty} x^2 dF(x) = c,$$

if

$$a^2(1-b)^2 + 2abd(1-b) + b^2c < c.$$

From this inequality it is seen that b lies between

$$1 \quad \text{and} \quad 1 - \frac{2(c - ad)}{a^2 - 2ad + c}$$

If the law of error is a Pearson type III,

$$y = k \left(1 + \frac{x}{q}\right)^{hq} e^{-hx},$$

then b lies between

$$1 \text{ and } 1 - \frac{2ke^{hq}(hq - ah + 2)\Gamma(hq + 1)}{a^2h^3(hq)^{hq} - 2akhe^{hq}\Gamma(hq + 1) + ke^{hq}(qh + 2)\Gamma(hq + 1)},$$

provided the denominator is not zero.

COROLLARY. *Under any law of error, whose second moment with respect to the true value a exists and whose first moment with respect to the true value is zero, there exist values of the constant b such that the probable value of the square of the error of bm is less than the probable value of the square of the error of m . These values satisfy the inequalities*

$$1 - \frac{2c}{a^2 + c} < b < 1.$$

The laws of error considered in the corollary include those which are symmetrical with respect to the true value and also others which are skew.

The preceding theorem does not apply to the law of error

$$f(x) = \frac{h}{\pi(1 + h^2x^2)},$$

for the second moment does not exist.

If the law of error for the error of m is

$$(3) \quad f(x) = \frac{2h}{\pi(1 + h^2x^2)^2},$$

then b satisfies the inequalities

$$1 - \frac{2}{a^2h^2 + 1} < b < 1.$$

If the error law is a Pearson type $X^\dagger$,

$$(4) \quad f(x) = \frac{h}{2} e^{-h|x|},$$

then b satisfies the inequalities

[†] Elderton, *Frequency Curves and Correlation*.

$$1 - \frac{4}{a^2 h^2 + 2} < b < 1.$$

If the error law is the Gaussian law,

$$\frac{h}{\sqrt{\pi}} e^{-h^2 x^2},$$

then the corollary to Theorem 1 becomes in this special case identical to a theorem given by Dodd.*

If the error law for the errors of the measurement m is $f(x) = 1/(2k)$ in $(-k, k)$ and zero elsewhere, then

$$1 - \frac{2k^2}{3a^2 + k^2} < b < 1.$$

This last law is used for the law of errors which are made in using a table of logarithms, for here the error is constant in a certain interval and zero elsewhere.

3. *Concerning the Error of the Mean.* If M is the mean and B is a constant it is desirable to know the nature of the probable value of certain functions of the error of BM and also the probable value of these same functions of the error of M . The following theorems will be useful in this respect.

THEOREM 2. *If the error law for the individual variable is of limited variation and symmetrical with respect to the true value, then the error law for the mean is also symmetrical with respect to the true value.*†

PROOF. Let $f(x)dx$ be the probability that the error lies in the interval $(x, x+dx)$. The probability that the sum lies in the interval $(u, u+du)$ is, according to a theorem given by Dodd,‡

$$\phi_n(u) = \frac{1}{\pi} \int_0^\infty g(t) \cos(ut) dt,$$

* Dodd, *The error-risk of certain functions of the measurements*, Monatshefte für Mathematik und Physik, vol. 24 (1913), pp. 268-276; Dodd, *The probability of the arithmetic mean compared with that of certain other functions of the measurements*, Annals of Mathematics, (2), vol. 14 (1912-13), pp. 186-198.

† Theorem 2 can be proved for the discrete case.

‡ Dodd, *The frequency law of a function of variables with given frequency laws*, Annals of Mathematics, (2), vol. 27 (1925-26), pp. 12-20.

where the parameter u does not enter the function $g(t)$. The function $p_n(u)$ is an even function and hence is symmetrical with respect to the origin which has been chosen as the true value. The law for the mean is obtained by replacing $p_n(u)$ by $p_n(nu)n$, and this is symmetrical to the true value.

In the theory of sampling from a parent population which is known, the standard error of the mean is equal to the standard error of the parent population divided by the square root of the number of the variates in the sample, provided the parent contains an infinite number of variates.* If the true value of the quantity to be measured is a and is considered to be the mean of the parent population the standard error of the mean can in general be found for the mean from that of the error law for the individual measurements. This idea is treated also in the theory of expected values leading up to the Tchebycheff inequalities.†

When the error laws for the individual errors of the measurements are the same, then the expected value of the mean of the errors is the same as the expected value of an individual error. This can be treated also by sampling theory.

THEOREM 3. *Let x be the error of the measurement m , d the expected value of x and c the expected value of x^2 . Then under any law of error whose second moment with respect to the true value a exists, there exist values of the constant B such that the probable value of the square of the error of BM is less than the probable value of the square of the error of M , provided*

$$adn - c \neq 0, \quad a^2n - 2adn + c \neq 0.$$

Under these conditions B lies between

$$1 \quad \text{and} \quad 1 - \frac{2(c - adn)}{a^2n - 2adn + c}.$$

The proof of this theorem is similar to that of Theorem 1.

COROLLARY. *Under any law of error, whose second moment exists and whose first moment with respect to the true value a is zero, there exist values of the constant B such that the probable value*

* C. H. Richardson, *The Statistics of Sampling*, a dissertation for the doctorate at The University of Michigan.

† Fisher, *The Mathematical Theory of Probability*, pp. 104-109.

of the square of the error of BM is less than the probable value of the square of the error of M ; these values satisfy the inequalities

$$1 - \frac{2c}{na^2 + c} < B < 1,$$

where M is the mean of n measurements and c is the expected value of the square of the error.

The proof of this corollary is similar to that of the Corollary to Theorem 1, since Theorem 2 shows that the error law for the mean is such that its first moment with respect to the true value is also zero.

If the law for the individual errors is (3), then the law for the mean of the errors is*

$$\frac{1}{\pi(hn)^n} \sum_{r=1}^{n+1} \binom{n}{r-1} \frac{\Gamma(r) \cos(r \cdot \tan^{-1} u)}{(hn)^{-n-1} (n^2 + n^2 h^2 u^2)^{r/2}}.$$

In this case the constant B satisfies the following inequalities:†

$$1 - \frac{2}{na^2 h^2 + 1} < B < 1,$$

where n measurements are made.

If the error law for the individual measurement is (4), then the law for the mean is‡

$$n h e^{-nh|u|} \sum_{r=0}^{n-1} \frac{n^{(-r)} h n |u|^{n-1-r}}{2^r (n-1-r)! (r)!}.$$

In this case

$$1 - \frac{4}{na^2 h^2 + 2} < B < 1.$$

If the error law for the errors of the measurement m is $f(x) = 1/(2k)$, then

* Baten, *Theorems concerning probability*, a dissertation for the doctorate at the University of Michigan, 1929. Karl Mayr, *Wahrscheinlichkeitsfunktionen und ihre Anwendungen*, Monatshefte für Mathematik und Physik, vol. 30 (1920), pp. 17-44.

† These inequalities for B can be obtained most readily from the three inequalities for b at the end of §2 by replacing h^2 by nh^2 ; and the validity of this process follows from the second paragraph above Theorem 3.

‡ See Baten, loc. cit.

$$1 - \frac{2k^2}{3na^2 + k^2} < B < 1.$$

If the error law is

$$f(x) = \frac{2h}{\pi(e^{hx} + e^{-hx})},$$

then

$$1 - \frac{2\pi^2}{4a^2h^2 + \pi^2} < b < 1,$$

and

$$1 - \frac{2\pi^2}{4na^2h^2 + \pi^2} < B < 1.$$

If the error function is

$$f(x) = \frac{ph}{2\Gamma(1/p)} e^{-h|p||x|^p},$$

we shall have

$$1 - \frac{2\Gamma(3/p)}{\Gamma(1/p) a^2h^2 + \Gamma(3/p)} < b < 1,$$

and

$$1 - \frac{2\Gamma(3/p)}{\Gamma(1/p) na^2h^2 + \Gamma(3/p)} < B < 1.$$

Other error laws may be used.

4. Concerning the Absolute Error of the Measurement.

THEOREM 4. *Under any law of error which is symmetrical with respect to the true value a , whose first moment with respect to the true value exists and which has only one maximum which is at the true value, there exist values of the constant b such that the probable value of the absolute error of bm is less than the probable value of the absolute error of the measurement m . A sufficient condition for such a value is*

$$1 - \frac{r}{ka^2 + r} < b < 1,$$

where r is one-half of the mean deviation and k is the maximum value, and a is assumed to be positive.

PROOF. The error $|x'| = |a - bm| = |a(1-b) + bx|$, where x is the error of the measurement m . The probable value of $|x'|$ is

$$P(|x'|) = \int_{-\infty}^{\infty} |a(1-b) + bx| f(x) dx,$$

where $f(x)$ is the error law. The quantity $\{a(1-b) + bx\}$ is negative when

$$x < -\frac{a(1-b)}{b} = -R.$$

Hence

$$\begin{aligned} P(|x'|) &= - \int_{-\infty}^{-R} \{a(1-b) + bx\} f(x) dx \\ &\quad + \int_{-R}^{\infty} \{a(1-b) + bx\} f(x) dx \\ &= a(1-b) \left\{ - \int_{-\infty}^{-R} f(x) dx + \int_{-R}^{\infty} f(x) dx \right\} \\ &\quad + a(1-b) \int_{-R}^R f(x) dx + b \left\{ - \int_{-\infty}^{-R} x f(x) dx \right. \\ &\quad \left. + \int_{-R}^{\infty} x f(x) dx \right\} + b \int_{-R}^R x f(x) dx \\ &= a(1-b) \int_{-R}^R f(x) dx + 2b \int_{-R}^{\infty} x f(x) dx \\ &< a(1-b) 2ka(1-b)/b + 2b \int_0^{\infty} x f(x) dx, \end{aligned}$$

which is less than

$$P(|x|) = \int_{-\infty}^{\infty} |x| f(x) dx,$$

if

$$2a^2k(1-b)^2/b + 2br < 2 \int_0^{\infty} x f(x) dx,$$

or if

$$2a^2k(1-b)^2 + 2b^2r - 2br < 0,$$

or if

$$1 - \frac{r}{a^2k + r} < b < 1.$$

If the error law is the Gaussian law then the above result is identical to that obtained by Dodd who used altogether the Gaussian law.

THEOREM 5. Under the law $f(x) = he^{-h|x|}/2$, there exist values of the constant b such that the probable value of the absolute error of bm is less than the probable value of the absolute error of the measurement m . A sufficient condition for such a value is

$$1 - \frac{1 - ah}{a^2h^2 - ah + 1} < b < 1,$$

where a is positive.

PROOF. By a similar process,

$$\begin{aligned} P(|x'|) &= ha(1-b) \int_0^R e^{-hx} dx + hb \int_R^\infty xe^{-hx} dx \\ &= a(1-b)(1 - e^{-hR}) + bRe^{-hR} + be^{-hR}/h. \end{aligned}$$

Since

$$e^{hR} = (1 + hR + h^2R^2/2 + \dots),$$

we have $e^{hR} > hR$ and $e^{-hR} < 1/(hR)$. We have also

$$\begin{aligned} 1 - e^{-hR} &= \left(1 - 1 + hR - \frac{h^2R^2}{2!} + \frac{h^3R^3}{3!} - \dots\right) \\ &= hR - \left(\frac{h^2R^2}{2!} - \frac{h^3R^3}{3!} + \dots\right). \end{aligned}$$

The last parenthesis is positive, therefore

$$1 - e^{-hR} < hR = a(1-b)h/b.$$

Therefore

$$P(|x'|) \leq ha^2(1-b)^2/b + hR + b/h,$$

which is less than $P(|x|) = 1/h$ if

$$h^2a^2 - 2a^2h^2b + a^2h^2b^2 + hba(1-b) + b^2 < b,$$

or

$$1 - \frac{1 - ah}{a^2h^2 - ah + 1} < b < 1.$$

The denominator is not zero unless ah , a , or h is imaginary. These cases are excluded here. If $1 > ah$ the fraction in the first inequality is positive. For other cases the inequalities have no meaning. Theorem 4 gives a lower limit for b regardless of the true value of a .

If the error law is

$$f(x) = \frac{2h}{\pi(1 + h^2x^2)^2},$$

we have

$$1 - \frac{1}{2a^2h^2 + 1} < b < 1.$$

Under the law

$$f(x) = \left(\frac{h}{4} + \frac{h^2|x|}{4} \right) e^{-h|x|},$$

we have

$$1 - \frac{3}{a^2h^2 + 3} < b < 1.$$

Under a type of Bayes function

$$f(x) = \frac{3h(1 - h^2x^2)}{4},$$

we have

$$1 - \frac{1}{4a^2h^2 + 1} < b < 1.$$

If

$$f(x) = \frac{2h}{\pi(e^{hx} + e^{-hx})},$$

we have

$$1 - \frac{\sum_0^{\infty} (-1)^n / (2n+1)^2}{ha^2 + \sum_0^{\infty} (-1)^n / (2n+1)^2} < b < 1.$$

Under the error law

$$f(x) = \frac{hp}{2\Gamma(1/p)} e^{-h|z|^p},$$

we have

$$1 - \frac{\Gamma(2/p)}{ph^2a^2 + \Gamma(2/p)} < b < 1.$$

Under the law $f(x) = 1/(2k)$, where k is a constant, we have

$$1 - \frac{k^2}{2a^2 + k^2} < b < 1.$$

5. Concerning the Absolute Value of the Error of the Mean.

If the Gaussian law is the error law for the individual errors, then the law for the mean of n is also another Gaussian law. According to Theorem 4, we may write

$$(5) \quad 1 - \frac{1}{2h^2a^2\sqrt{n+1}} < B < 1,$$

where B is a constant which is multiplied by the mean M . This indicates that the mean deviation of the mean is larger than the mean deviation of B times the mean, provided the law of error is the Gaussian law. This indicates that there is a better value than the mean for the true value, if the law is the Gaussian law.

The method of random sampling from a known parent population may be employed to investigate the nature of B for other laws of error. Let the parent population consist of the variates $x_1, x_2, x_3, \dots, x_s$. Let samples of r variates be taken from this parent. It has been shown[†] that the mean of all possible sample means is equal to the mean of the parent population, if repetitions of samples are not allowed.

Let the first sample mean of r variates be

$$Z_1 = \frac{x_1 + x_2 + x_3 + \dots + x_r}{r},$$

then the deviation of the first sample mean from the mean of the sample means is, in absolute value,

$$\begin{aligned} |\bar{Z}_1| &= \left| \frac{x_1 + x_2 + x_3 + \dots + x_r}{r} - M_x \right| \\ &= \left| \frac{(x_1 - M_x) + (x_2 - M_x) + (x_3 - M_x) + \dots + (x_r - M_x)}{r} \right| \\ &= \left| \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \dots + \bar{x}_r}{r} \right| \\ &\leq \frac{|\bar{x}_1| + |\bar{x}_2| + |\bar{x}_3| + \dots + |\bar{x}_r|}{r}, \quad \bar{x}_i = x_i - M_x, \end{aligned}$$

where the mean of the parent population is M_x .

[†] An editorial, *Annals of Mathematical Statistics*, vol. 1 (1930), pp. 101-121. See also the first footnote on page 61.

The mean deviation of the sample means is found by finding the average of all possible $|\bar{Z}|$'s and this is clearly less than or equal to the mean deviation of the parent population. The mean deviation of the sample means is

$$\frac{\sum_{i=1}^{sC_r} |\bar{Z}_i|}{sC_r},$$

while the mean deviation of B times the sample means is

$$\frac{\sum_{i=1}^{sC_r} B |\bar{Z}_i|}{sC_r} = \frac{B \sum_{i=1}^{sC_r} |\bar{Z}_i|}{sC_r},$$

where the summations are divided by sC_r since there are that many possible samples that can be taken from the s variates. The mean deviation of B times the sample means is less than the mean deviation of the sample means if B is less than unity. If s goes to infinity the above is also true.

The above inequalities for the Gaussian law give more information than the above sampling method. B is not only less than 1 but it lies in a certain interval with the end points deleted. This is found in (5). Going back to sampling, if the mean of the parent population is the true value then the mean deviation of B times the sample means is less than the mean deviation of the mean of the sample means.

To be able to find the inequalities for B when other laws of error are used it is necessary to find the law for the mean when the law for the individual error law is known.

The limits for b are not always the best limits, for under Theorem 1 the lower limit for b is found by setting the first derivative of (2) equal to zero and solving for b . This gives the middle of the interval as the lower limit of b , which brings b nearer to unity.

Intervals for the constant b can be found when the higher moments are used.

In general the coefficients b and B are so near to unity that bm and BM do not differ appreciably from m and M respectively, yet this is not always the case, notably, when the standard error of m is large relative to m .

CHAPTER III

COMPARISON OF CERTAIN PROBABILITIES *

Probabilities of a continuous variable may be considered as areas under a probability curve between certain limits. The entire area under the curve for the interval of definition is unity, which represents the sum of all possible probabilities of the variable under consideration. The probability that the variable x lies on the interval (a, b) is $\int_a^b f(x)dx$, where $f(x)dx$ is the probability that x lies on the interval $(x, x+dx)$.

When the probability function or probability law for the variable x is known, the probability laws for functions of x can be determined.¹ In some cases it is necessary to compare probabilities of various functions of x . These comparisons are obtained by considering two areas under the respective probability laws on an interval of comparison. Generally this interval of comparison is a small interval which includes the abscissa of the greatest value of the law. In error laws the interval which is of most importance is that which includes the probable value, or the mean.

In the following theorems several intervals of comparison will be considered. Frequently the theorems are only true in the interval mentioned.

THEOREM 1. *A necessary and sufficient condition that the probability of bx is greater than the probability of x , is that the positive constant b is less than 1; provided the interval of comparison is $(-c, d)$ and that x extends over the interval $(-k, h)$, where $k > c$ and $h > d$.*

¹ If the date (400 A.D.) assigned to Vyāsa by S. N. Dās Gupta be correct, the date of invention of the modern notation in India would be pushed back several centuries earlier. But in that case it would be difficult to explain the presence of the enunciation of the notation in the *Āryabhaṭīyam*.

² D. E. Smith, *History of Mathematics*, Vol. II (1925), pp. 64 and 65.

³ Ibid, p. 72. Also Kaye, *Indian Mathematics*, p. 31.

⁴ Smith and Karpinski, *The Hindu-Arabic Numerals*, pp. 4 and 10. Also Cajori, *A History of Mathematics*, (1922), pp. 102 and 103.

⁵ See Dodd: *The frequency law of a function of one variable*, Bulletin of the American Mathematical Society, Vol. 31 (1925), pp. 27-31.

* Published in The American Mathematical Monthly, May, 1932

PROOF: Let $f(x)dx$ be the probability that the variable x lies on the interval $(x, x+dx)$, then the probability that, $y=bx$, lies on the interval $(y, y+dy)$ is¹

$$F(y)dy = (1/b)f(y/b)dy.$$

If the probability that y lies on the interval $(-c, d)$ is greater than the probability that x lies on this interval, then

$$(1/b) \int_{-c}^d f(y/b)dy > \int_{-c}^d f(x)dx,$$

or

$$\int_{-c/b}^{d/b} f(z)dz > \int_{-c}^d f(x)dx.$$

If this is true the interval $(-c/b, d/b)$ is greater than the interval $(-c, d)$, or

$$(1/b)(d+c) > (d+c),$$

or

$$b < 1.$$

The sufficiency of the condition can be proved by reversing the steps of the foregoing proof.

It is readily seen that the quantity c must be non-negative.

The quantity k may be zero, in which case c is zero. If the maximum of the law is not at the origin, the origin can be changed to the maximum or greatest value of the probability function.

If each variable x_i ($i=1, 2, 3, \dots, n$) is subject to the same probability law defined on the interval $(-k, k)$, then the law for the sum of n variables can be found.² Let $y=(1/n)\sum_{i=1}^n x_i$, then according to the above theorem the probability of the mean for the interval $(-c, d)$ is greater than the probability for the sum, since here $b=1/n$, and is less than unity when n is greater than 1. The probability law for the mean may be zero in part of $(-c, d)$.

THEOREM 2. *Let the probability law for the variable x be $f(x)$, which is defined on the interval $(-c, a)$, where c and a are positive quantities. If the positive quantity b is less than 1, there exists an interval of comparison $(0, d)$ in which the probability of bx^k is greater than the probability of x^s , provided $k > s > 1$.*

PROOF: If the law for x is $f(x)$, then the law for $y=bx^k$ is³

$$F(y) = \frac{p \cdot y^{(1-k)/k}}{kb^{1/k}} \cdot f\{(y/b)^{1/k}\},$$

and the law for $z=x^s$ is

$$\Psi(z) = \frac{q \cdot z^{(1-s)/s}}{s} \cdot f\{(z)^{1/s}\},$$

¹ Dodd, loc. cit., Bulletin.

² Dodd, *The frequency law of a function of variables with given frequency laws*, Annals of Mathematics, Ser. 2, vol. 27 (1925) pp. 12-20.

³ Dodd, loc. cit., Bulletin.

where p and q are 2 if k and s are even and 1 if k and s are odd. The probability that bx^k lies on the interval $(0, d)$ is greater than the probability that x^s lies on this interval if

$$\int_0^d F(y)dy > \int_0^d \Psi(z)dz, \text{ or if}$$

$$\frac{p}{kb^{1/k}} \int_0^d v^{(1-k)/k} f\{(y/b)^{1/k}\} dy > \int_0^d (q/s) z^{(1-s)/s} f\{(z)^{1/s}\} dz,$$

or if

$$\int_0^{(d/b)^{1/k}} f(t)dt > \int_0^{d^{1/s}} f(v)dv.$$

Let $s = rk$, where $r < 1$. The last inequality is true if

$$(d/b)^{1/k} > (d)^{1/rk},$$

or if

$$\log d < \{r/(r-1)\} \log b.$$

The quantity $r/(r-1)$ is negative and the logarithm of b is negative; hence the logarithm of d is less than some positive number, which determines a limit for the interval of comparison $(0, d)$.

The theorem is true if $s=k$ as can be seen from

$$(d/b)^{1/k} > (d)^{1/k}.$$

The theorem is also true if $s=1$.

THEOREM 3. *If $\sum_{i=1}^n b_i < \sum_{i=1}^m k_i$, then the probability of $\sum_{i=1}^n b_i x_i$ is greater than the probability of $\sum_{i=1}^m k_i x_i$, provided the probability law for each variable is*

$$f(x_i) = \frac{h}{\pi(h^2 + x_i^2)},$$

and the interval of comparison is $(-c, d)$.

PROOF: The probability function for $y = b_i x_i$, is

$$F(y) = \frac{b_i h}{\pi} \cdot \frac{1}{b_i^2 h^2 + y^2},$$

the probability function for $\sum_{i=1}^m k_i x_i = u$, is

$$\Psi(u) = \frac{\sum_{i=1}^m k_i h}{\pi} \cdot \frac{1}{h^2 \left(\sum_{i=1}^m k_i \right)^2 + u^2},$$

the probability function for $w = \sum_{i=1}^n b_i x_i$, is

$$\theta(w) = \frac{\sum_{i=1}^n b_i x_i}{\pi \left(h^2 \left(\sum_{i=1}^n b_i \right)^2 + w^2 \right)}$$

The probability that the sum $\sum_{i=1}^n b_i x_i$ lies in the interval $(-c, d)$ is greater than the probability that the sum $\sum_{i=1}^m k_i x_i$ lies in this interval if

$$\frac{h \sum_{i=1}^n b_i}{\pi} \int_{-c}^d \frac{du}{h^2 \left(\sum_{i=1}^n b_i \right)^2 + u^2} > \frac{h \sum_{i=1}^m k_i}{\pi} \int_{-c}^d \frac{dw}{h^2 \left(\sum_{i=1}^m k_i \right)^2 + w^2},$$

and this is true since $\sum_{i=1}^n b_i < \sum_{i=1}^m k_i$.

Nothing was stated about each b_i and each k_i being less than 1. In fact the theorem is true if some are greater than 1 and some less than 1, as long as the sum of the b 's is less than the sum of the k 's. If $k_i = 1$ then the probability of $\sum_{i=1}^n b_i x_i$ is greater, for the interval $(-c, d)$, than the probability of $\sum_{i=1}^n x_i$, if $\sum_{i=1}^n b_i < n$.

THEOREM 4. *If the probability function for the variable x_i , ($i = 1, 2, \dots, n$), is the Gaussian function, then the probability that*

$$-c \leq \sum_{i=1}^n b_i x_i \leq d,$$

is greater than the probability that

$$-c \leq \sum_{i=1}^m k_i x_i \leq d,$$

provided

$$\sum_{i=1}^n b_i^2 < \sum_{i=1}^m k_i^2.$$

PROOF: If the law for x_i is $h/\sqrt{\pi} \cdot e^{-h^2 x_i^2}$, then according to a method given by Dodd¹ the laws for the sums $\sum_{i=1}^n b_i x_i = u$ and $\sum_{i=1}^m k_i x_i = w$ are respectively

$$F(u) = \frac{h}{(\pi)^{1/2} \left(\sum_{i=1}^n b_i^2 \right)^{1/2}} e^{-h^2 u^2 / \sum_{i=1}^n b_i^2},$$

$$\Psi(w) = \frac{h}{(\pi)^{1/2} \left(\sum_{i=1}^m k_i^2 \right)^{1/2}} e^{-h^2 w^2 / \sum_{i=1}^m k_i^2}.$$

¹ Dodd, loc. cit., Annals.

The theorem is true if

$$\int_{-c}^d F(u) du > \int_{-c}^d \Psi(w) dw,$$

which is true if

$$\sum_{i=1}^n b_i^2 < \sum_{i=1}^m k_i^2.$$

This shows that the probability of the sum $\sum_{i=1}^m x_i$ is greater than the probability of the sum $\sum_{i=1}^n x_i$, on the interval $(-c, d)$, if $m < n$.

In a weighted mean the sum $\sum b_i = 1$, hence the arithmetic mean has the greater probability for the interval mentioned. If however the $\sum_{i=1}^n b_i^2$ is less than $1/n$, then this kind of weighted mean has greater probability than the probability of the mean.

If the law for the individual variable x_i is an even function the probability for the sum $\sum_{i=1}^n b_i x_i$ can be compared with the probability for the sum $\sum_{i=1}^m k_i x_i$ by examining two integrals of the types¹

$$p_{nb}(u) = (1/\pi) \int_0^\infty f(b_1 t) f(b_2 t) \cdots f(b_n t) dt,$$

$$p_{mk}(u) = (1/\pi) \int_0^\infty f(k_1 t) f(k_2 t) \cdots f(k_m t) dt.$$

The nature of the functions under the integral signs will have to be examined in each particular case before it can be stated that the probability of the first is greater than the probability of the second for some interval of comparison.

THEOREM 5. *Under the Gaussian Law of error the probable value of any given power of the absolute error of one weighted mean with weights $p_1, p_2, \dots, p_n$, is less than that of a second weighted mean with weights $q_1, q_2, \dots, q_m$, if $\sum_{i=1}^n p_i^2 < \sum_{i=1}^m q_i^2$.*

PROOF: If the law for the individual variables is $he^{-h^2 x^2}/\sqrt{\pi}$, then the law for the first weighted mean is

$$P_n(u) = \frac{P}{(\pi)^{1/2}} e^{-P^2 u^2}, \text{ where } P^2 = h^2 / \sum_{i=1}^n p_i^2;$$

and the law for the second weighted mean is

$$Q_m(w) = \frac{Q}{(\pi)^{1/2}} e^{-Q^2 w^2}, \text{ where } Q^2 = h^2 / \sum_{i=1}^m q_i^2.$$

¹ Dodd, loc. cit., *Annals*.

The probable value of the absolute error of the first raised to the power k is less than the absolute error of the second raised to this same power if

$$\int_{-\infty}^{\infty} |u|^k P_n(u) du < \int_{-\infty}^{\infty} |w|^k Q_m(w) dw,$$

or if

$$1/P^2 < 1/Q^2, \text{ or if } \sum_{i=1}^n p_i^2 < \sum_{j=1}^m q_j^2.$$

This shows that under this error law the probable value of any given positive power of the absolute error of the arithmetic mean is less than that of any other weighted mean. In this case $m=n$ and the weights are each equal to $1/n$.

If $p_i=1/m$ and $q_j=1/n$, where $i=1, 2, \dots, m$ and $j=1, 2, \dots, n$, then the probable value of any given positive power of the absolute error of the first weighted mean, (here the real mean), is less than that of the second mean if $m < n$. This shows that the mean with a greater number of variates is the better. This is as it should be, or what would be expected. But the p 's and q 's can be such that even if $m > n$ the probable value of any given positive power of the absolute error of the first weighted mean is greater than that of the second weighted mean. The weighted mean with the most variates is not as good as that of the second with a smaller number of variates.

CHAPTER IV

THE EVALUATION OF CERTAIN DEFINITE
INTEGRALS BY THE USE OF
PROBABILITY FUNCTIONS*

1. *Introduction.* The object of this paper is to present three methods of evaluating certain definite integrals by using probability functions. The first method consists in finding the probability law or function for the sum of n independent variables, which are each subject to given probability laws, by Mayr's method† and comparing this with the probability for the same sum obtained by Dodd's method.‡ On equating the two results thus obtained one often finds the value of certain integrals.

Method II consists in finding the probability law for n independent variables, which is expressed as an integral, and then allowing n to be equal to 1.

† Karl Mayr, *Wahrscheinlichkeitsfunktionen und ihre Anwendungen* Monatshefte für Mathematik und Physik, vol. 30.

‡ E. L. Dodd, *The frequency law of a function of one variable*, this Bulletin, vol. 31 (1925), pp. 27–31; *The frequency law of a function of variables with given frequency laws*, Annals of Mathematics, (2), vol. 27, pp. 12–20.

* Presented to the American Mathematical Society, August 29, 1929. Published in the Bulletin of the American Mathematical Society, June, 1930.

Method III is founded on setting up a probability function containing the integral to be evaluated as a factor, determining a second probability function defined over the same interval, and then finding a third probability function for a chosen relation of the variables under consideration. The integral need not be a factor of the first function, but should arise somewhere in the process.

The first two methods, as far as the writer knows, seem to be new.

2. *Method I.* Consider the integral

$$(1) \quad \int_0^{\infty} \frac{\cos(ut - nk)dt}{(q^2 + t^2)^{nr/2}},$$

where $k = r \tan^{-1}(t/q)$, and $r > -1$. Let x_i be taken arbitrarily out of the interval $(0, \infty)$ and let

$$\frac{q^r}{\Gamma(r)} x_i^{r-1} e^{-qx_i} dx_i$$

be the probability that x_i lies in the interval $(x_i, x_i + dx_i)$. The probability function for the sum $x_1 + x_2 + x_3 + \dots + x_n = u$ can be written in two different forms following Mayr's or Dodd's method. These will be denoted hereafter by $P_n(u)$ and $F_n(u)$, respectively. We may then write, omitting du ,

$$P_n(u) = \frac{q^{nr} u^{nr-1} e^{-qu}}{(nr-1)!},$$

$$F_n(u) = \frac{q^{nr}}{\pi} \int_0^{\infty} \frac{\cos(ut - nk)dt}{(q^2 + t^2)^{nr/2}}.$$

Equating these, we find

$$\int_0^{\infty} \frac{\cos(ut - nk)dt}{(q^2 + t^2)^{nr/2}} = \frac{\pi u^{nr-1} e^{-qu}}{(nr-1)!}.$$

Evaluating the integral, we obtain

$$(2) \quad \int_{-\infty}^{\infty} e^{-h^2 x(x-u)} dx.$$

Let x_i be taken arbitrarily out of the interval $(-\infty, \infty)$ and let the probability function for x_i be

$$\frac{h}{\sqrt{\pi}} e^{-h^2 x_i^2} dx_i.$$

The probability function for the sum $x_1 + x_2 = u$ is

$$P_2(u) = \frac{h^2}{\pi} \int_{-\infty}^{\infty} e^{-h^2 u^2} e^{-2h^2 x(x-u)} dx, \quad (\text{Mayr's method}),$$

$$F_2(u) = \frac{h}{(2\pi)^{1/2}} e^{-h^2 u^2/2}, \quad (\text{Dodd's method}).$$

Hence we may write

$$\int_{-\infty}^{\infty} e^{-2h^2 x(x-u)} dx = \frac{\sqrt{\pi}}{h\sqrt{2}} e^{h^2 u^2/2}.$$

Let x_i be taken arbitrarily out of the interval $(-\infty, \infty)$ and let the probability function for x_i be

$$\frac{2h}{\pi} \frac{1}{e^{hx_i} + e^{-hx_i}}.$$

The probability function for the sum $x_1 + x_2 = u$ is

$$P_2(u) = \frac{4h^2}{\pi^2} \int_{-\infty}^{\infty} \frac{e^{hu} e^{hx} dx}{(e^{2hu} + e^{2hx})(e^{2hx} + 1)},$$

$$F_2(u) = \frac{4h^2}{\pi} \int_0^{\infty} \frac{\cos ut \, dt}{(e^{\pi t/(2h)} + e^{-\pi t/(2h)})^2} = \frac{4h^2 u}{\pi^2 (e^{hu} - e^{-hu})};$$

hence

$$(3) \quad \int_{-\infty}^{\infty} \frac{e^{hx} dx}{(e^{2hu} + e^{2hx})(e^{2hx} + 1)} = \frac{u}{e^{2hu} - 1}.$$

Let the probability function for x_i be $1/k$ for $-k/2 \leq x_i \leq k/2$, and zero elsewhere. The probability function for the sum $x_1 + x_2 = u$, by Mayr's method, is

$$P_2(u) = \begin{cases} 0, & \text{for } (-\infty < u \leq -k), \\ \frac{u+k}{k^2}, & \text{for } (-k \leq u \leq 0), \\ -\frac{u+k}{k^2}, & \text{for } (0 \leq u \leq k), \\ 0, & \text{for } (k \leq u < \infty); \end{cases}$$

and by Dodd's method it is

$$F_2(u) = \frac{4}{k^2\pi} \int_0^\infty \frac{\sin^2(kt/2) \cos(ut) dt}{t^2}.$$

Hence we have

$$(4) \quad \frac{4}{k^2\pi} \int_0^\infty \frac{\sin^2(kt/2) \cos(ut) dt}{t^2} = \begin{cases} 0, & \text{for } (-\infty < u \leq -k), \\ \frac{u+k}{k^2}, & \text{for } (-k \leq u \leq 0), \\ -\frac{u+k}{k^2}, & \text{for } (0 \leq u \leq k), \\ 0, & \text{for } (k \leq u < \infty). \end{cases}$$

In a similar manner, by using the probability law stated above for x_i , and the probability law for the sum $x_1+x_2+x_3$, or $x_1+x_2+x_3+x_4$, we get

$$\frac{6}{k^3\pi} \int_0^\infty \frac{\sin^3(kt/2) \cos(ut) dt}{t^3} = \begin{cases} 0, & \text{for } \left(-\infty < u \leq -\frac{3k}{2}\right), \\ \frac{1}{2k^3} \left(u + \frac{3k}{2}\right)^2, & \text{for } \left(-\frac{3k}{2} \leq u \leq -\frac{k}{2}\right), \\ \frac{1}{4k^3} (-4u^2 + 3k^2), & \text{for } \left(-\frac{k}{2} \leq u \leq \frac{k}{2}\right), \\ \frac{1}{2k^3} \left(-u + \frac{3k}{2}\right)^2, & \text{for } \left(\frac{k}{2} \leq u \leq \frac{3k}{2}\right), \\ 0, & \text{for } \left(\frac{3k}{2} \leq u < \infty\right); \end{cases}$$

$$\frac{2^4}{k^4\pi} \int_0^\infty \frac{\sin^4(kt/2) \cos(ut) dt}{t^4}$$

$$= \begin{cases} 0, & \text{for } (-\infty < u \leq -2k), \\ \frac{1}{6k^4}(u+2k)^3, & \text{for } (-2k \leq u \leq -k), \\ \frac{1}{6k^4}(-3u^3 - 6ku^2 + 4k^3), & \text{for } (-k \leq u \leq 0), \\ \frac{1}{6k^4}(3u^3 - 6ku^2 + 4k^3), & \text{for } (0 \leq u \leq k), \\ \frac{1}{6k^4}(-u+2k)^3, & \text{for } (k \leq u \leq 2k), \\ 0, & \text{for } (2k \leq u < \infty). \end{cases}$$

By continuing this process, a probability function for the sum of n independent variables is obtained, which is continuous and has different analytic expressions in different intervals of its existence, yet which can be expressed as a single definite integral.

3. *Method II.* If $f_i(x_i) = 3(1-x_i^2)/4$ is the probability law for x_i for $-1 \leq x_i \leq 1$, and zero elsewhere, then the probability law for the sum $x_1 + x_2 + x_3 + \dots + x_n = u$, is

$$(5) \quad F_n(u) = \frac{3^n}{\pi} \int_0^\infty \frac{(\sin t - t \cos t) \cos(ut) dt}{t^{3n}}.$$

For $n=1$, this integral is evidently $f_i(u)$. Hence

$$(6) \quad \int_0^\infty \frac{(\sin t - t \cos t) \cos(ut) dt}{t^3} = \frac{\pi}{4}(1-u^2), \quad (-1 \leq u \leq 1).$$

In particular, for $u=0, 1$, we have

$$(7) \quad \int_0^\infty \frac{(\sin t - t \cos t) dt}{t^3} = \frac{\pi}{4},$$

$$(8) \quad \int_0^\infty \frac{(\sin t - t \cos t) \cos t dt}{t^3} = 0.$$

If the probability function for x_i is $1/(2k)$, $-k \leq x_i \leq k$, and zero elsewhere, where k is a certain constant, then the law for the sum of n independent variables is, according to Dodd's method,

$$(9) \quad F_n(u) = \frac{1}{\pi} \int_0^\infty \frac{(\sin kt)^n \cos(ut) dt}{k^n t^n}.$$

If we set $n=1$ and $u=0$, we find the well known integrals

$$(10) \quad \int_0^\infty \frac{\sin(kt) \cos(ut) dt}{t} = \frac{\pi}{2}, \quad (-k \leq u \leq k),$$

$$(11) \quad \int_0^\infty \frac{\sin(kt) dt}{t} = \frac{\pi}{2}, \quad (k > 0).$$

Let the probability law for x_i be $(\cos x_i)/2$ for $-\pi/2 \leq x_i \leq \pi/2$, and zero elsewhere; the law for the sum of n variables is

$$(12) \quad F_n(u) = \frac{1}{\pi} \int_0^\infty \frac{\cos^n(t\pi/2) \cos(ut) dt}{(1-t^2)^n}.$$

If we set $n=1$, this becomes

$$(13) \quad \int_0^\infty \frac{\cos(t\pi/2) \cos(ut) dt}{(1-t^2)} = \frac{\pi}{2} \cos u, \quad (-\pi/2 \leq u \leq \pi/2).$$

In particular, with $u=0$,

$$(14) \quad \int_0^\infty \frac{\cos(t\pi/2) dt}{(1-t^2)} = \frac{\pi}{2}.$$

If the probability law for x_i is

$$\frac{h}{\sqrt{\pi}} e^{-h^2 x_i^2},$$

for $(-\infty < x_i < \infty)$, then the law for the sum of the squares, $x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2 = u$, is

$$(15) \quad F_n(u) = \frac{h^n}{\pi} \int_0^\infty \frac{\cos(ut - nm) dt}{(h^4 + t^2)^{1/4}}, \quad \tan 2m = \frac{t}{h^2}.$$

For $n=1$, we find

$$(16) \quad \frac{h}{\pi} \int_0^\infty \frac{\cos(ut - m) dt}{(h^4 + t^2)^{1/4}} = \frac{h e^{-h^2 u}}{(\pi u)^{1/2}}.$$

When written in full, this becomes

$$\frac{h}{2\pi} \int_0^\infty \frac{[(h^4 + t^2)^{1/2} + h^2]^{1/2} \cos(ut) - [(h^4 + t^2)^{1/2} - h^2]^{1/2} \sin(ut) dt}{(h^4 + t^2)^{1/4}} \\ = \frac{he^{-h^2 u}}{(\pi u)^{1/2}}.$$

4. *Method III.* We shall now show how to evaluate the integral

$$(17) \quad J = \int_0^\infty \frac{e^{-px} \sin(qx) dx}{x}.$$

Let

$$f(x) = \frac{1}{J} \cdot \frac{e^{-px} \sin qx}{x}, \quad g(y) = e^{-y},$$

respectively, be the probability functions for the variables x and y for $0 \leq x, y < \infty$, and let $y = xs$; then the probability function for s is*

$$P(s) = \frac{1}{J} \int_0^\infty \frac{e^{-px} \sin(qx) e^{-sx} dx}{x} = \frac{1}{J} \cdot \frac{q}{q^2 + (p + s)^2}.$$

Since $P(s)$ is a probability function for the interval $(0, \infty)$, we have $\int_0^\infty P(s) ds = 1$. Therefore

$$\frac{1}{J} \int_0^\infty \frac{q ds}{q^2 + (p + s)^2} = \frac{1}{J} \left(\frac{\pi}{2} - \tan^{-1} \frac{p}{q} \right) = \frac{1}{J} \left(\tan^{-1} \frac{q}{p} \right).$$

Hence

$$J = \int_0^\infty \frac{e^{-px} \sin(qx) dx}{x} = \tan^{-1} \frac{q}{p}.$$

Consider the integral

$$(18) \quad \int_0^\infty \frac{ds}{1 + s^4}.$$

Let

$$f(x) = \frac{\sqrt{2} \sin x}{\pi \sqrt{x}}, \quad g(y) = \frac{2}{\sqrt{\pi}} e^{-y^2},$$

where x and y lie in the interval $(0, \infty)$, and let $y = s\sqrt{x}$. Then we have

* See the second footnote on p. 35

$$(19) \quad P(s) = \frac{2\sqrt{2}}{\pi} \int_0^\infty \frac{e^{-xs^2} \sqrt{x} \sin x \, dx}{\sqrt{x}} = \frac{2\sqrt{2}}{\pi(1+s^4)}.$$

Since we know that $\int_0^\infty P(s) ds = 1$, we may write

$$\int_0^\infty \frac{ds}{1+s^4} = \frac{\pi}{2\sqrt{2}}.$$

We shall now prove that

$$(20) \quad \int_0^\infty \frac{\cos x \, dx}{x+a} = \int_0^\infty \frac{xe^{-ax} dx}{1+x^2}.$$

Let

$$f(x) = \frac{1}{J} \cdot \frac{\cos x}{x+a}, \quad g(y) = e^{-y},$$

for $0 \leq x, y < \infty$, where J is the integral on the left of the above equality. The probability law for s , where $y = (x+a)s$, is

$$(21) \quad P(s) = \frac{1}{J} \int_0^\infty \frac{\cos x \, e^{-(x+a)s} (x+a) dx}{(x+a)} = \frac{1}{J} \cdot \frac{se^{-as}}{1+s^2}.$$

But $\int_0^\infty P(s) ds = 1$. Hence

$$\frac{1}{J} \int_0^\infty \frac{se^{-as} ds}{1+s^2} = 1,$$

or

$$(22) \quad \int_0^\infty \frac{se^{-as} ds}{1+s^2} = J = \int_0^\infty \frac{\cos s ds}{s+a},$$

which agrees essentially with (20).

CHAPTER V

SIMULTANEOUS TREATMENT OF DISCRETE AND
CONTINUOUS PROBABILITY BY USE
OF STIELTJES INTEGRALS *

The object of this paper is to present several theorems pertaining to the probability that certain functions lie within certain intervals. The first theorem is a generalization of Markoff's Lemma, which is proven for the discrete and continuous cases by use of the accumulative frequency function and Stieltjes integrals. Tchebycheff's Theorem is obtained as a corollary to a very general theorem, the proof of which is based upon the first theorem. Other corollaries are given.

Three theorems, due to Guldberg, which follow are concerned with the probability that a non-negative chance variable be less than certain functions of the expected value of the variable. These are proved for the discrete and continuous cases by employing accumulative frequency functions and Stieltjes integrals. This is the first time, as far as the writer knows, the discrete and continuous cases for these theorems have been included in a single proof.

Theorem 1. If A denotes the expected value of the non-negative variable x and t is any number greater than 1, then the probability that $x \leq At^2$ is greater than $1 - \frac{1}{t^2}$.

Proof: If x is a discrete variable with values at x_i ($i = 1, 2, \dots, n$) with corresponding probabilities p_i , then it is understood that the probability that x takes other values is zero. If x is a continuous variable having a probability function defined over the interval (a, b) , then it is understood that the probability that x lies outside of (a, b) is zero in case (a, b) is different from $(-\infty, +\infty)$. In both cases x is a continuous variable in the interval $(-\infty, +\infty)$. Let the probability that x lies in the interval $(-\infty, x)$ be $F(x)$, with $F(-\infty) = 0$ and $F(+\infty) = 1$. Then the probability that x lies in the interval (x_1, x_2) is

$$F(x_2) - F(x_1) + \frac{1}{2} \left\{ F(x_2 + 0) - F(x_2 - 0) \right\} + \frac{1}{2} \left\{ F(x_1 + 0) - F(x_1 - 0) \right\}$$

* Published in The Annals of Mathematical Statistics, February, 1930.

where the last two limits are different from zero when there is probability different from zero at x_1 and x_2 . This exists since $F(x)$ is a non-decreasing function over the interval $(-\infty, +\infty)$. In the special case when $F(x) = \int_{-\infty}^x f(x) dx$ where $f(x)$ is summable, $f(x) dx$ represents the probability that x lies in the interval $(x, x+dx)$.

In either case, by definition

$$A = \int_{-\infty}^{\infty} x \cdot dF(x)$$

$x > At^2$ in the interval $(At^2 + \epsilon, \infty)$, where ϵ approaches 0, hence

$$A > \int_{At^2 + \epsilon}^{\infty} x \cdot dF(x)$$

But

$$\int_{At^2 + \epsilon}^{\infty} x dF(x) = \lim_{\epsilon \rightarrow 0} \int_{At^2 + \epsilon}^{\infty} x dF(x) = \lim_{\epsilon \rightarrow 0} z_{\epsilon} \int_{At^2 + \epsilon}^{\infty} dF(x) = \lim_{\epsilon \rightarrow 0} z_{\epsilon} \cdot \lim_{\epsilon \rightarrow 0} \int_{At^2 + \epsilon}^{\infty} dF(x)$$

by the first theorem of the mean, which holds for Stieltjes integrals in this case. Here $z > At^2 + \epsilon$, hence $\lim_{\epsilon \rightarrow 0} z_{\epsilon} \geq At^2$, therefore $A > At^2 \int_{At^2 + \epsilon}^{\infty} dF(x)$ But $\int_{At^2 + \epsilon}^{\infty} dF(x)$ is the probability P that x is greater than At^2 , hence

$$A > At^2 P, \quad Q > 1 - \frac{1}{t^2}$$

where Q is the probability that $x \leq At^2$.

This theorem is a generalization of Markoff's Lemma¹, which he proved for the discrete case. The above proof takes care of the discrete case, the continuous case and the case which is a combination of the discrete and continuous.

Theorem 2. If $f(x_1, x_2, \dots, x_n)$ is a function of n independent variables, then the probability that

$$|f - k| \leq t \sqrt{E(f^2) - 2kE(f) + k^2}$$

1. "Wahrscheinlichkeitsrechnung," by Markoff. 1912. Page 54.

is greater than $1 - 1/t^2$; where E represents the expected value, k is a constant and $t > 1$.

Proof: Let

$$y = \{f(x_1, x_2, \dots, x_n) - k\}^2 \quad \text{then}$$

$$E(y) = E(f^2) - 2kE(f) + k^2$$

By theorem 1 the probability that

$$|f - k| \leq t \sqrt{E(f^2) - 2kE(f) + k^2}$$

is greater than $1 - 1/t^2$.

Corollary: If $f(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n$ and $k = \sum_{i=1}^n E(x_i)$, then theorem 2 becomes the famous Tchebycheff theorem¹. This theorem is: If $x_1, x_2, \dots, x_n$ be n independent variables, then the probability that

$$\left| \sum_{i=1}^n x_i - \sum_{i=1}^n E(x_i) \right| \leq t \sqrt{\sum_{i=1}^n E(x_i^2) - 2 \sum_{\substack{i,j=1 \\ i \neq j}}^n E(x_i x_j) + \left\{ \sum_{i=1}^n E(x_i) \right\}^2},$$

is greater than $1 - 1/t^2$.

This proof is by far simpler than that given by Tchebycheff, while it is similar to that given by Markoff.

In the corollary if $k = \sum_{i=1}^n E(x_i)$, $E(x_i) = a$, $E(x_i^2) = A$, then the probability that

$$\left| \frac{\sum x_i}{n} - a \right| \leq \frac{t}{\sqrt{n}} \sqrt{A - a^2}$$

is greater than $1 - 1/t^2$.

If $f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i^s$ then the probability that

$$\left| \sum_{i=1}^n x_i^s - k \right| \leq t \sqrt{\sum_{i=1}^n A_{i,2s} + 2 \sum_{i=1}^n a_{i,s} a_{i,s} - 2k \sum_{i=1}^n a_{i,s} + k^2}$$

is greater than $1 - 1/t^2$, where the variables are independent, $E(x_i^{2s}) = A_{i,2s}$; $E(x_i^s) = a_{i,s}$. If s is negative it is under-

1. "Des Valeurs Moyennes," by Tchebecheff, Journal de Math. 1867 (2). Vol. 12.

stood that x can not take on the value zero, if $s = a/b$, where a is odd and b is even, it is understood that x is non-negative.

Other interesting results may be obtained from this theorem if $f(x_1, x_2, \dots, x_n)$ represents various functions of the n independent variables and k be given different values. If f is the sum of the variables, theorem 2 is a more general theorem than Tchebycheff's theorem because of the constant k which may have values other than $\sum_{i=1}^n E(x_i)$.

Let x_i be the result of an individual throw of a coin, $x_i = 1$ if a head is thrown and $x_i = 0$ if a tail is thrown; then $E(x_i^2) = p \cdot 1 + q \cdot 0$, where p is the probability of a head and q is the probability of a tail. Let m represent the number of heads thrown in n throws and let $k = np \pm \sqrt{n - npq}$, then the probability that

$$|m - (np \pm \sqrt{n - npq})| \leq t\sqrt{n}, \text{ or that}$$

$$\left| \frac{m}{n} - \left(p \pm \sqrt{\frac{1}{n} - \frac{pq}{n}} \right) \right| \leq \frac{t}{\sqrt{n}}, \text{ is greater than } 1 - \frac{1}{t^2}.$$

Let $t = \frac{1}{\sqrt{n}}$, then the probability that

$$-\frac{1}{\sqrt{n}} \pm \sqrt{\frac{1}{n} - \frac{pq}{n}} \leq \frac{m}{n} - p \leq \frac{1}{\sqrt{n}} \pm \sqrt{\frac{1}{n} - \frac{pq}{n}}$$

is greater than $1 - \frac{1}{t^2}$ or $1 - \frac{1}{\sqrt{n}}$, which approaches unity as n increases. It is near unity for large values of n . This shows that the empirical probability approaches the true probability p as the number of throws increases, and the advantage of k .

Theorem 3: Let $u'_{n,x}$ be the expected value of the non-negative variable x raised to the power n and t any number greater than 1, then the probability that $x \leq t^n \sqrt[n]{u'_{n,x}}$ is greater than $1 - \frac{1}{t^n}$.

Proof: Let $c > \sqrt[n]{u'_{n,x}}$, and let $F(x)$ be the probability that x lies in the interval $(-\infty, x)$, then by definition

$$u'_{n,x} = \int_{-\infty}^{\infty} x^n dF(x); \text{ and } \frac{u'_{n,x}}{c^n} = \int_{-\infty}^{\infty} x^n dF(x)/c^n$$

Now

$$\begin{aligned} \frac{u'_{n,x}}{c^n} &> \int_{\frac{c}{t}}^{\infty} x^n dF(x) / c^n = \lim_{\epsilon \rightarrow 0} \int_{c+\epsilon}^{\infty} x^n dF(x) / c^n \\ &= \lim_{\epsilon \rightarrow 0} \frac{(ze)^n}{c^n} \cdot \lim_{\epsilon \rightarrow 0} \int_{c+\epsilon}^{\infty} dF(x) > i \cdot \int_{c+\epsilon}^{\infty} dF(x), \end{aligned}$$

by the first theorem of the mean, and since $\lim_{\epsilon \rightarrow 0} \frac{(ze)^n}{c^n} \geq 1$.

Since $\int_{c+\epsilon}^{\infty} dF(x)$ is the probability P that x is greater than c ,

$$\frac{u'_{n,x}}{c^n} > p \quad \text{or} \quad Q > 1 - \frac{u_{n,x}}{c^n}$$

where Q is the probability that $x \leq c$.

Let $t = \frac{c}{\sqrt[n]{u'_{n,x}}}$, then $\frac{1}{t^n} = \frac{u'_{n,x}}{c^n}$, hence

$$Q > 1 - \frac{1}{t^n}.$$

But Q becomes the probability that $x \leq \sqrt[n]{u'_{n,x}}$, since c was any number greater than $\sqrt[n]{u'_{n,x}}$.

Let $y = |x - k|$, then theorem 3 becomes: If $u'_{n,y}$ is the expected value of $|x - k|^n$ and t is greater than 1, then the probability that $|x - k|$ does not surpass the multiple $t \sqrt[n]{u'_{n,y}}$, is greater than $1 - \frac{1}{t^n}$, where k is a constant.

If $k = \int_{-\infty}^{\infty} x \cdot dF(x)$, then $u'_{n,y}$ becomes $u'_{n,\bar{x}}$ and theorem 3 states that the difference $|x - k|$ does not surpass the multiple $t \sqrt[n]{u'_{n,\bar{x}}}$, is greater than $1 - \frac{1}{t^n}$. In this special case theorem 3 becomes Guldberg's theorem¹, but this is more general than his theorem, for it includes the continuous case, the discrete case and the case which is a combination of the discrete and continuous.

If $y = |f(x) - k|$ is used for the variable, a more general theorem is obtained. Here $f(x)$ is a function of x . Of course, the probability law for $f(x)$ must be secure from that of x if the continuous case is under consideration. Certain restrictions must be placed upon $f(x)$ concerning continuity, summability and concerning the inverse.

1. "Sur un théorème de M. Markoff," by Alf. Guldberg. *Compte Rendue*, Vol. 175. (1922) page 679.

Theorem 4. The probability that the difference $|x - m|$ is not greater than the multiple $t \alpha_{r,x}$, $t > 1$, is greater than $1 - \left(\frac{\alpha_{r,x}}{\sqrt{t-1}} \right)^n$ ($t > 1$), where $\alpha_{r,x}$ is the expected value of $|x - m|^r$, and m is the expected value of x .

Theorem 5. The probability that the positive quantity x does not surpass the multiple $t m$, ($t > 1$), is greater than $1 - \left(\frac{\alpha_{r,x}}{\sqrt{t-1}} \right)^n$ ($t > 1$), where $\alpha_{r,x}$ is the expected value of $|x - m|^r$, and $m = E(x)$.

These last two theorems are due to Guldberg¹ for the discrete case. By the method used in theorem 3 these can be proven for the continuous case, the discrete case, and the case which is a combination of the discrete and continuous.

1. "Sur quelques inégalités des le calcul de probabilités," by Guldberg, *Comp. Rend.* Vol. 175 (1922), p. 1382.
"Sur le théorème de Tchebecheff," by Guldberg, *Comptes Rendue*, Vol. 175 (1922), p. 418.

The author wishes to thank Professor Harry C. Carver for encouragement and helpful criticism given in the beginning and throughout the preparation of this paper; also to acknowledge indebtedness to Professors James W. Glover, George Y. Rainich and James A. Shohat for reading the first draft and offering helpful suggestions.

Copified by
The Coplfyer Corporation
Detroit, Michigan